

EKSTRAKSI FITUR BERBASIS AVERAGE FACE UNTUK PENGENALAN EKSPRESI WAJAH

Jarot Dwi Prasetyo¹⁾, Zaehol Fatah²⁾, Taufik Saleh³⁾

¹ Manajemen Informatika, AMIK Ibrahimy Situbondo
email: jarot_dwi_prasetyo@yahoo.com

² Manajemen Informatika, AMIK Ibrahimy Situbondo
email: zaeholfatah@gmail.com

³ Manajemen Informatika, AMIK Ibrahimy Situbondo
email: tausho501@gmail.com

Abstract

In recent years it appears interest in the interaction between humans and computers. Facial expressions play a fundamental role in social interaction with other humans. In two human communications is only 7% of communication due to language linguistic message, 38% due to paralanguage, while 55% through facial expressions. Therefore, to facilitate human machine interface more friendly on multimedia products, the facial expression recognition on interface very helpful in interacting comfort. One of the steps that affect the facial expression recognition is the accuracy in facial feature extraction. Several approaches to facial expression recognition in its extraction does not consider the dimensions of the data as input features of machine learning

Through this research proposes a wavelet algorithm used to reduce the dimension of data features. Data features are then classified using SVM-multiclass machine learning to determine the difference of six facial expressions are anger, hatred, fear of happy, sad, and surprised Jaffe found in the database. Generating classification obtained 81.42% of the 208 sample data.

Keywords: wavelet, SVM, facial expression recognition, human computer interaction

1. PENDAHULUAN

Dalam tahun-tahun terakhir ini terjadi ketertarikan dalam interaksi antara manusia dan komputer. Ekspresi wajah memainkan peranan yang mendasar dalam interaksi sosial dengan manusia lainnya. Dalam komunikasi dua orang manusia hanya 7% pesan komunikasi disampaikan lewat bahasa linguistic, 38% paralanguage, sedangkan 55% ditransfer melalui ekspresi wajah^[15]. Oleh karena itu, untuk memfasilitasi antarmuka manusia-mesin yang lebih bersahabat pada produk multimedia, analisis ekspresi wajah berbasis vision banyak diteliti pada sepuluh tahun terakhir.

Ekman dan Friesen^[4] mengembangkan sistem yang paling komprehensif untuk sintesa ekspresi wajah berbasis apa yang mereka sebut Action Units (AU). Mereka mendefinisikan sistem pengkodean aksi wajah (Facial Action Coding System (FACS)). FACS terdiri atas 46 AU, yang menggambarkan gerakan wajah dasar. Yang berbasis pada aktivitas otot dan menggambarkan secara detail pengaruh masing-masing AU pada fitur wajah. Mereka mendefinisikan enam emosi dasar yaitu bahagia, sedih, ketakutan, jijik, terkejut dan marah. Masing-masing emosi dasar ini

memiliki ekspresi wajah yang unik. Di awal tahun 1990 komunitas engineering mulai menggunakan hasil-hasil ini untuk membuat metode otomatis untuk pengenalan emosi dari ekspresi wajah dalam gambar atau video^[22].

Essa dan Pentland^[12] menyajikan hasil pengenalan dan pengklasifikasian emosional tunggal dari ekspresi wajah berbasis metode aliran optik yang dihubungkan dengan geometris, fisik dan model wajah berbasis gerakan. Mereka menggunakan energi gerakan 2 dimensi dan template yang mengkodekan keduanya yaitu magnitude dan arah gerakan. Dengan mempelajari gerakan 2 dimensi 'ideal' untuk empat ekspresi emosional (marah, jijik, bahagia dan terkejut), mereka mendefinisikan spatial-temporal template untuk ekspresi-ekspresi tersebut. Walaupun pendekatan yang diusulkan oleh Essa dan Pentland belum sepenuhnya divalidasi, perlu dicatat bahwa spatial-temporal template ekspresi wajah dari metode unik untuk klasifikasi dari ekspresi wajah^[9], berhasil untuk pengenalan real time, pengklasifikasian emosional dan membuat emosi untuk ekspresi emosional dasar. Ini diwujudkan pada robot wajah 3 dimensi yang dapat mengenali dan menghasilkan ekspresi

emosional. Mereka menggunakan data distribusi kecerahan citra wajah dan neural network back-propagation 3 lapis untuk klasifikasi dan sintesis ekspresi wajah.

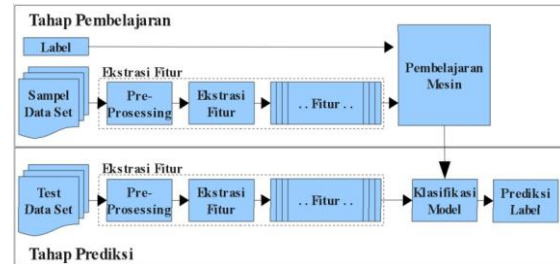
Pada paper ^[16] mengusulkan metode pengenalan emosi menggunakan model Naive Bayes (NB). Mereka mengusulkan kerangka kerja untuk memilih distribusi model untuk masing-masing emosi. Hasil eksperimen menunjukkan bahwa penggunaan asumsi model Cauchy memberikan akurasi klasifikasi yang lebih baik daripada asumsi model Gaussian.

Pada paper ^[20] mengusulkan pengenalan wajah berbasis ciri citra wajah dari filter Gabor dengan pengklasifikasian menggunakan Support Vector Machine (SVM). Untuk memperkecil dimensi vector ciri dari 163.840 yang dihasilkan dari filter Gabor, Weimin menggunakan Principle Component Analysis (PCA), sehingga diperoleh 106 komponen dasar. Dari hasil penelitiannya, keakurasian pengenalan mencapai lebih dari 80% bila jumlah data pelatihan lebih dari 70. Penggunaan PCA memberikan keakurasian yang lebih tinggi dan signifikan bila jumlah data pelatihannya besar.

Pengenalan ekspresi wajah sebagian besar tidak mengambil perubahan gerak pada fitur wajah^[8]. Beberapa penelitian memanfaatkan gerak fitur pada wajah berbasis video. Dengan mengambil fitur geometris pada wajah yang dicari (bentuk vektor ^[2], parameter animasi wajah ^[17], sudut dan jarak ^[10], lintasan ^[19]), atau perbedaan antara permukaan wajah dalam urutan frame (seperti optical flow ^[14], dan differential-AAM ^[21], atau tekstur dan perubahan gerak dalam wilayah local wajah (seperti perubahan bentuk permukaan ^[6], satuan gerak ^[11], spatiotemporal descriptors ^[7], satuan animasi ^[5], dan perbedaan pixel ^[11]). Meskipun mendapatkan hasil yang baik, pendekatan ini membutuhkan ketepatan lokasi dan mengikuti titik pada wajah, seperti yang di jelaskan pada paper ^[13]. Selain itu, masih dalam pertanyaan bagaimana belajar tata bahasa dalam mendefinisikan ambiguitas dalam data input ^[18]. Dalam kata lain, teknik pengenalan ekspresi wajah berdasarkan gambar adalah sebuah alternatif lain untuk mengenali emosi berdasarkan penampilan dalam sebuah gambar, dan kondisi yang penting ketika hanya beberapa gambar yang tersedia untuk pembelajaran dan pengujian.

2. HASIL DAN PEMBAHASAN

Gambar 1 mengilustrasikan bagaimana cara kerja pengenalan ekspresi wajah, dengan menggabungkan pemrosesan, pembelajaran dan kondisi prediksi. Pada kondisi sebelum diproses, dengan mendeteksi citra wajah.



Gambar 1. Langkah Kerja

2.1 Data Set dan Representasi

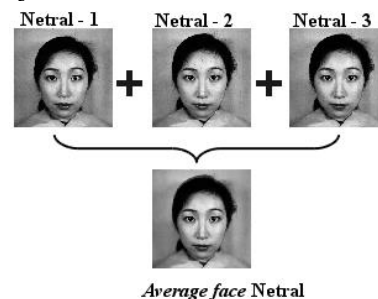
Data base yang digunakan dalam proses percobaan berjumlah 213 citra ekspresi wajah perempuan jepang. Jumlah citra yang sesuai untuk masing-masing dari 7 kategori ekspresi (netral, senang, sedih, terkejut, marah, jijik dan takut) kira-kira sama. Beberapa dari mereka yang ditunjukkan pada gambar 2. untuk rincian tentang koleksi citra ini, disebut dalam ^[3].



Gambar 2. Contoh Database Ekspresi Wajah (JAFPE)

2.2 Ekstraksi Wajah

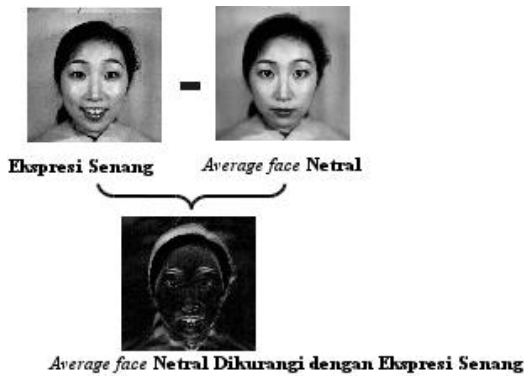
Ekstraksi fitur bertujuan untuk mendapatkan fitur dari citra, teknik untuk ekstraksi fitur yang digunakan adalah menggunakan metode average face yaitu menghitung rata-rata dari citra wajah netral.



Gambar 3. Average Face dari Wajah Netral

Pada gambar 3 adalah proses untuk mencari nilai rata-rata dari ekspresi wajah netral dengan menjumlahkan piksel pada posisi yang sama kemudian dirata-ratakan.

Setelah mendapatkan nilai rata-rata dari ekspresi wajah netral, kemudian dari masing-masing ekspresi dikurangi dari rata-rata ekspresi netral.



Gambar 4. Pengurangan antara Average Face dengan ekspresi

Gambar 4 adalah proses pengurangan antara citra average face netral dengan ekspresinya. Setelah dikurangi maka ROI bagian wajah dengan ukuran 125x106, sehingga yang diambil pada area wajah.

Hasil dari proses pengurangan wajah terlihat pada Gambar 5. Hasil dari ROI ini yang akan dimasukkan kedalam mesin pembelajaran.



Gambar 5. ROI wajah 125x106

2.3 Klasifikasi dengan SVM

Penggunaan SVM pada prinsipnya menyelesaikan sebuah permasalahan quadratic programming. Oleh karena itu, proses pelatihannya hampir sama dan tahapannya dapat dilihat pada gambar 1, Akan tetapi untuk pembelajaran tak-terbimbing dengan SVM data pelatihan dan data pengujian adalah data yang sama. Selain itu, untuk proses pelatihannya dapat juga hanya menggunakan sebagian data dari data pengujian sehingga proses waktu

pelatihan menjadi lebih singkat, tetapi hal ini mungkin menurunkan akurasi pada tahap pengujian. Sedangkan untuk pembelajaran terbimbing justru sebaliknya, dapat meningkatkan akurasi pada tahap pengujiannya. Dalam penelitian ini menggunakan SVM dengan kernel RBF (SVM-RBF) dan SVM-LINEAR yang terdapat pada weka. Tipe SVM yang digunakan dalam pembelajaran mesin adalah nu-SVC, dalam prosesnya nu-SVC digunakan untuk data yang besar dengan skema pembelajaran one-vs-one dengan beberapa parameter yang digunakan adalah C : 1.0; Kernel : rbf; degree : 3; Gamma : 0.0; coef0 : 0.0; Probability : True; shrinking : True; tol : 1e-3; cache_size : 400; class_weight : {dict, 'auto'}; Verbose : False; max_iter : -1.

Dengan strategi pengujian cross-validasi, dalam k-kali cross-validasi, sampel asli secara acak dibagi menjadi sub-samples k. Dari subsamples k, suatu subsampel tunggal dipertahankan sebagai validasi data untuk pengujian model, dan k tersisa -1 subsamples digunakan sebagai data pelatihan. Proses cross-validasi kemudian diulang kali k (lipatan), dengan masing-masing subsamples k digunakan tepat sekali sebagai data validasi. Hasil k dari kemudian dapat dirata-ratakan (atau dikombinasi) untuk menghasilkan estimasi.

2.4 Hasil

Dimensi fitur yang dihasilkan dari average face adalah 125x106, yaitu sebesar 13.251 atribut. dengan metode pembelajaran menggunakan SVM-RBF maka menghasilkan keakurasian sebesar 90.31% seperti terlihat pada tabel 1.

Tabel 1. Hasil Uji Coba

Prediksi Aktual	Marah	Benci	Takut	Senang	Sedih	Terkejut	Total
Marah	29	1	0	0	0	0	96.67%
Benci	0	27	2	0	0	0	93.10%
Takut	1	2	26	0	2	1	81.25%
Senang	0	0	0	28	3	0	90.32%
Sedih	0	1	3	1	26	0	83.87%
Terkejut	0	0	1	0	0	29	96.67%
RATA-RATA							90.31%

3. KESIMPULAN

Penelitian ini menyajikan tentang klasifikasi ekspresi wajah, dengan menggunakan database JAFFE (Japanese Female Facial Expression) sebagai sampel data, terdiri dari 213 citra

ekspresi wajah wanita dari sepuluh wanita yang masing-masing terdiri dari tujuh kategori ekspresi yaitu netral, bahagia, sedih, terkejut, marah, jijik dan takut. Dengan citra asli telah dipotong dan diskala ulang sedemikian rupa sehingga pada posisi mata memiliki jarak yang sama, dengan resolusi citra 256 piksel dan 256 piksel pada skala keabu-abuan.

Tahap berikutnya adalah ekstraksi fitur, beberapa metode ekstraksi fitur yang dipakai dalam penelitian ini diantaranya adalah average face. Data hasil ekstraksi fitur ini yang nantinya digunakan sebagai masukan pada mesin pembelajaran. Mesin pembelajaran menggunakan metode SVM dengan menghasilkan tingkat keakurasian 90.31% diperoleh dari data fitur dengan metode average face.

4. REFERENSI

- [1] W. Burleson A. Kapoor and R.M. Picard. 2007. "Automatic Prediction of Frustration". *Int'l J. Human-Computer Studies*, 65:724–736.
- [2] T. Yan C. Jixu And L Qiang. 2010. "A Unified Probabilistic Framework For Spontaneous Facial Action Modeling And Under Standing". *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32(2):258-273.
- [3] P. Ekman. 1990. "An Argument For Basic Emotions". *Cognition and Emotion*, 6(3):315–331.
- [4] P. Ekman And W. V. Friesen. 1978. "Facial Action Coding System: Investigator's Guide". Palo Alto, CA: Consulting Psychologists Press.
- [5] F. Dornaika and F. Davoine. 2008. "Simultaneous Facial Action Tracking And Expression Recognition In The Presence Of Head Motion". *Int'l J. Computer Vision*, 76:257–281.
- [6] F. Tsalakanidou and S. Malassiotis. 2010. "Real Time 2D+3D Facial Action And Expression Recognition". *Pattern Recognition*, 43:1763–1775.
- [7] G. Zhao and M. Pietikainen. 2009. "Boosted Multi -Resolution Spatio Temporal Descriptors For Facial Expression Recognition". *Pattern Recognition letters*, 30:1117–1127.
- [8] Z. ZN M. P Antic G.I. Roisman and T.S. Huang. 2009. "A Survey Of Affect Recognition Methods: A Udio, Visual, And Spontaneous Expressions". *IEEE Trans. Pattern Analysis and Machine Intelligence*, 3L(1):39 – 58.
- [9] H. Kobayashi and F. Hara. 1997. "Facial Interaction Between Animated 3D Face Robot And Human Beings". *SMC '97*, p. 3732 – 3737.
- [10] S.A. Ali Hamdi, V. Roberto and G. Theo. 2010. "Eyes Do Not Lie : S Pontaneous Versus Posed Smiles". *Proc. Int'I Conf. Multimedia*.
- [11] A. Garg L.S. Chen I. Cohen, N. Sebe and T.S. Huang. 2009. "Facial Expression Recognition From Video Sequence : T Emporal And Static Modeling". *Computer Vision and Image Understanding*, 91:160–187.
- [12] I.A. Essa and A.P. Pentland. 1997. "Coding Analysis Interpretation And Recognition Of Facial Expressions". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 (7) :757 – 763.
- [13] L. Peng and S.J.D. Pritce. 2009. "Toint And Implicit Registration For Face Recognition". *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, p. 1510–1517.
- [14] B. Bulot M. Yeasin And R. Sharma. 2006. "Recognition Of Facial Expressions And Measurement Of Levels Of Interest From Video". *IEEE Trans. Multimedia*, 8(3):500–508.
- [15] T. Mehrabian. 1968. "Communication Without Words". *Psychology Today*. Volume 2.
- [16] A. Garg M. Lew N. Sebe, I. Cohen and T. Huang. 2002. "Emotion Recognition Using A Cauchy Naive Bayes Classifier". *ICPR 2002*, 4:3732 – 3737.
- [17] P.S. A Leksic and A.K. Katsaggelos . 2006. "Automatic Facial Expression Recognition Using Facial Animation Parameters And Multistream Hmms". *IEEE Trans. Information Forensics and Security*, 1(1):3–11.
- [18] M. Pantic T. Huang, A. Nijholt and A. Pentland. 2007. "Human Computing And Machine Understanding Of Human Behavior : A Survey". *Artificial Intelligence for Human Computing*, 4451:47–71.
- [19] T.H. Wang and J.F.I. Lien. 2009. "Facial Expression Recognition System Based On Rigid And Non - Rigid Motion Separation

- And 3D Pose Estimation”. J. Pattern Recognition, 42(5):917–962.
- [20] X. Weimin. 2006. “Facial Expression Recognition Based On Gabor Filter And SVM”. Chinese Journal of Electronics, 15(4A).
- [21] Y. Cheon and D. Kim. 2009. “Natural Facial Expression Recognition Using Differential Aam And Manifold Learning”. Pattern Recognition, 42:1340–1350.
- [22] Y. Yacoob and L.S. Davis. 1996. “Recognizing Human Facial Expressions From Long Image Sequences Using Optical Flow”. IEEE Trans. Patt. Anal. Machine Intell., 18(6):636 – 642.